EMOTION DETECTION BY ELECTROENCEPHALOGRAM

by:

Jagruti Patel

(111EC0182)

Under the able guidance of

Prof. Debiprasad Priyabrata Acharya

A thesis presented for the degree of

Bachelor of Technology

Department of Electronics and Communication Engineering

National Institute of Technology, Rourkela

Certificate

This is to certify that the report titled ***EMOTION DETECTION BY ELECTROENCEPHALOGRAM***, submitted by Jagruti Patel, of National Institute of Technology, Rourkela for the partial fulfilment of the requirements for the degree of Bachelor of Technology, is a bonafide record of the work done by her in the department of Electronics and Communication Engineering under my supervision.

Prof. Debiprasad Priyabrata Acharya,
Department of Electronics &
Communication Engineering,
NIT Rourkela.

Acknowledgements

The enthusiastic guidance and support of Prof. Debiprasad Priyabrata Acharya inspired me to stretch beyond the limits. His profound insight has guided my thinking to improve upon the final prototype of the product. My solemnest gratefulness to him. I am also obliged to Queen Marry University, UK for providing me the data for the emotion classification and feature extraction. I am very much thankful to my family and all my friend for being supportive through out.

Jagruti Patel
111EC0182
NIT Rourkela

Abstract

Emotion Recognition is probably one of the biggest challenges engineers are facing now. With ongoing development and demand of man - machine interface it has been very necessary to implement emotion recognition. But emotion not being standardized or quantifiable, it is very difficult to classify. Again the lack of proper benchmark for differentiating between various emotions makes it a more difficult challenge. Again the outcome of emotions can be noticed in EEG signals - which again are very hard to classify. According to biomedical science, the EEG data can be classified in not amplitude domain, but in frequency domain - thus making the work more challenging.

This project uses machine learning with other statistical variations like PCA to find the benchmark for emotions of persons having similar behavioural characteristics and classify the emotions. The first stage of the work was detection of Epilepsy and classify the signals into epileptic and non-epileptic. This task hes been achieved by using conventional statistical moments like mean and variance. On a later stage machine learning was applied to classify.

The later part of work was to use nonlinear SVM and get emotion data for training and classification. A number of emotions like valence, arousal, dominance etc. Further the non-linear classification algorithm was tested on several dataset. The final stage of work includes HDL coding for implemetation of non-linear SVM.

Keywords: *EEG, Emotion detection, Epilepsy, SVM, Classification, PCA*

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

The electrical activity of brain in animals was first described and analysed by German psychiatrist Hans Berger. He later introduced the term Electroencephalogram, which is the recorded electrical acitivity in brain along the scalp. The method of recording these signals is called Electroencephalography, commonly known as EEG. EEG is a non-invasive method and widely used for recording and analysis of brain anomalies, traumas, psychological disease detection, research on human cognitive activities and many more. The emotion centre of the brain generates a huge lot of EEG signals and they vary according to emotions. The most important thing about the EEG signals are that, they are not consistent in amplitude domain and the variation is too much to classify them in terms of their amplitude. But study shows significant difference in frequency domain in different types of states and brain activities. For example when a person is asleep we can observe a particular frequency dominating and a different frequency dominates when the person is doing some heavy thinking. So, often frequency domain classification of EEG signal is adapted. The challenge of this project is to observe the pattern of those variations and try to classify the emotions, so that the goal of designing a perfect man machine interface can be achieved.

1.1 Study of EEG Signals

EEG measures voltage fluctuations resulting from ionic current that flows within the neurons of the brain. Though the EEG signal is not periodic, the features lies in the frequency domain analysis. Table I gives the details of the several classes of the EEG signals. The database for analysis were collected from [2]

Electric recordings from the exposed surface of the brain or from the outer surface of the head demonstrate continuous oscillating electric activity within the brain. Both the intensity and the patterns of this electric activity are determined to a great extent by the overall excitation of the brain resulting from functions in the brainstem reticular activating system (RAS). The undulation in the

Band	**Frequency**
Delta	$<$ 4Hz
Theta	4 - 7Hz
Alpha	8 - 15Hz
Beta	16 - 31Hz
Gamma	30 - 100Hz

Table 1.1: List of various classes of EEG signals

recorded electric potentials are called brain waves and the entire record is called electroencephalogram. Figure 1.1 shows the various brain waves and figure 1.2 shows the brain waves during various types of epilepsy

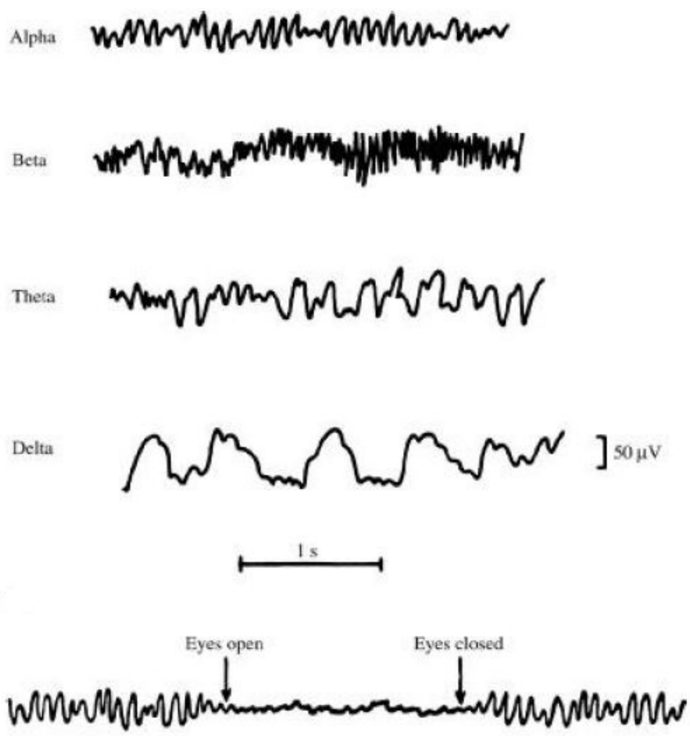

Figure 1.1: Different forms of brain waves including alpha, beta, theta, delta and change of alpha rythm when subject opens eyes

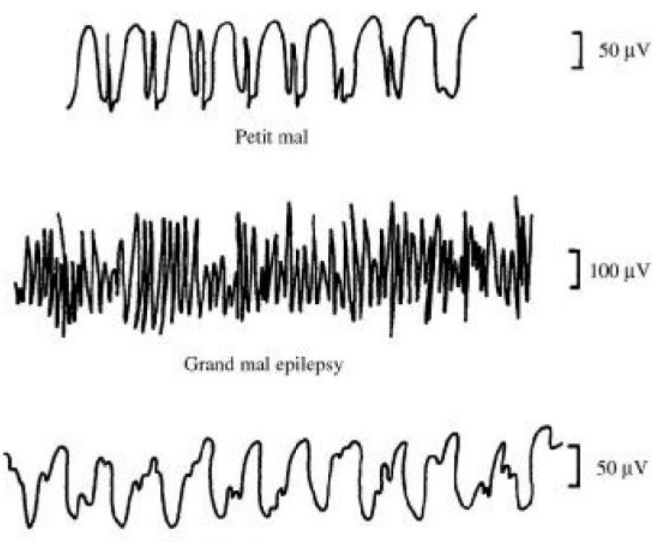

Figure 1.2: Various types of epileptic brain wave signals

Chapter 2

EPILEPSY DETECTION USING CLASSICAL METHODS

EEG signals are most prominently used in medical science for epilepsy detection. Thus the first stage of the work was to collect EEG data - both epileptic and non-epileptic. Further the method of simple first order and second order statistical comparison between the data was adapted for epilepsy detection.

2.1 Classical Method

Epilepsy is a group of neurological disorders characterized by epileptic seizures. It can often be confirmed with an electroencephalogram (EEG) but a normal test does not rule out the disease [2],[3]. Thus it is very necessary to predict the nature and the out-come accurately. Classification using simple statistical moments like mean, variance is the simplest method of analysing the EEG signals. A database of epilepsy was collected from [5]. The data represents the EEG signal of a person over time. Some of the signals contain seizure activity. A classical method of finding the population mean and population variance and comparing them to a threshold was applied first, but the results were catastrophic. Following may be the reasons of the failure.

- Variable gain of different channel
- Noise effect
- Non consistent data even in the time of no epilepsy.
- Situational spikes in the data
- Extremely large numbers of observations of the sample.

From basic statistics it is clear that if the data set is too large and there are anomalies in the data it could change the mean position and the output is not as expected. Thus in machine learning it is always advised to ignore the exceptions and extremes. Further study of the data and brain activity suggests that epilepsy usually affect the gamma waves [2], and the data is needed to be properly filtered.

2.2 Classical Method with Improvisation

Large sampling theory, or testing of hypothesis suggests that, when dealing with a very large data set, the sample data collection should be such that the sample contains maximum data of our interest. Thus to study the epileptic seizure, our data sample must be the particular class that is maximum affected by the seizure. As depicted earlier, we need to analyse the gamma signals of the brain wave only, so that any other type of interference (like emotional overwhelming) may not affect our data. Therefore a filter was designed to filter out the gamma waves. As gamma waves also contains the powerline frequency i.e. 50Hz, it is advised to filter the 50Hz artefacts from the data by using a *notch filter* [1]. Later a bandpass filter was designed to filter out the signals ranging from 32Hz to 100Hz - the Gamma signal.

For designing the band pass filter, several options were considered. Different types of Hamming, Hann and Kaiser windows were designed and implemented. Among them Kaiser window gave the best performance, thus the Kaiser window was preferred for designing the bandpass filter. The Properties of the designed Kaiser filter is listed below, and the response diagram is given in figure 2.1.

- Sampling frequency = 256Hz
- Order = 500
- Fc1 = 32
- Fc2 = 100
- beta = 20

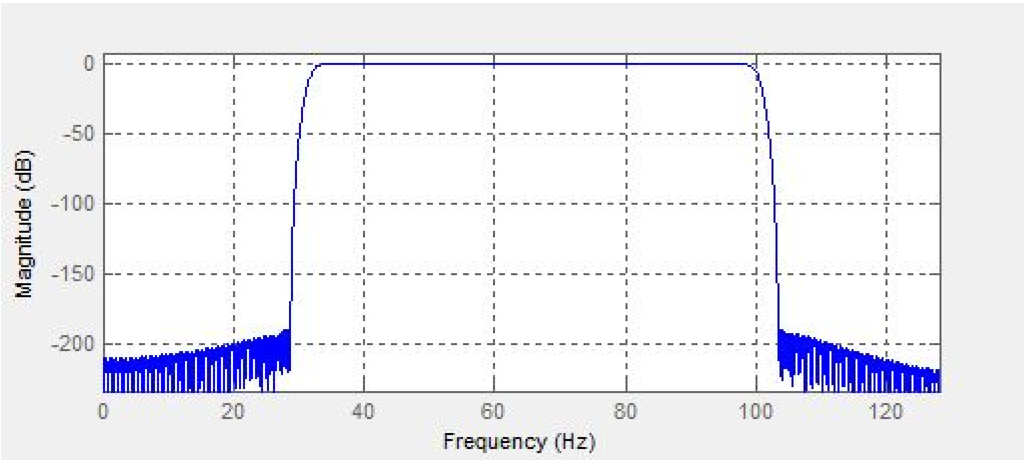

Figure 2.1: Response of Kaiser filter

The samples were then transformed into frequency domain by using fast Fourier transformation. It is clear from the response that the signals having frequency below 32Hz or above 100Hz were clipped. As the sample size is now reduced, and there is very few external interference, second order moment was applied to separate the anomalous data. Thus a threshold was chosen and taking the second order moment of the data as a feature, the data was classified. This gave significant improvement over the previous results. The figure 2.2 gives a sample result of the classical method.

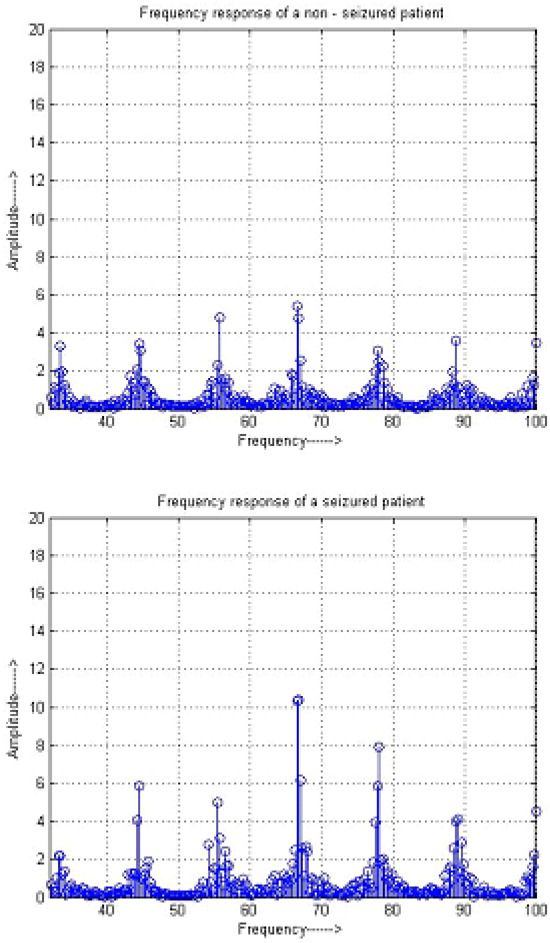

Figure 2.2: Comparison of results of seizure and no seizure data in classical method

2.3 Failure of Classical Method

After keen analysis it was found that the algorithm does not perform well when the data is in the boundary of seizure and no seizure activity (figure 2.3) and hence the method is not of much use when it comes to multiple features.Basically the performance degrades at the boarder of the classifier, which can be overcome by using support vector machine.

2.4 Principal Component Analysis

Principal component analysis (PCA) is a technique widely used for comparison or classification of large dimensional data. The technique is to reduce the dimensionality of the data by eliminating correlated data. By eliminating the correlation, the program can find the orthogonal members of the data which are mathematically uncorrelated, thus by presenting the exact number of features present in the data. In a nut shell PCA has an inbuilt ability to project the higher dimensional data into a lower dimension which helps us for the data redundancy part and increases the computation speed [1]. The main feature of PCA is that the user does not have to extract features, but the process itself separates the data which are mutually orthogonal or in technical term mutually independent data. This helps

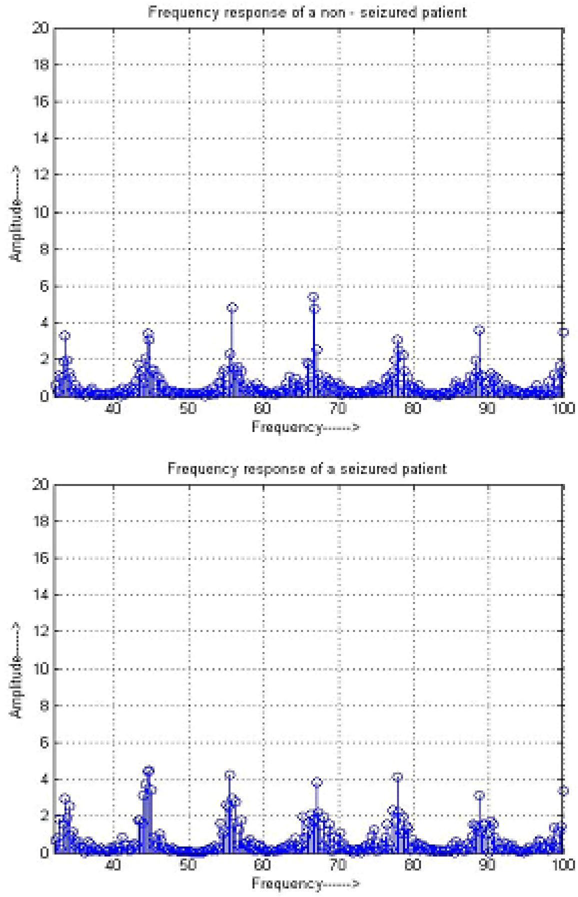

Figure 2.3: Comparison of results of seizure and no seizure data in classical method at boundary

for the classification of the data without being worried about the features and extraction of features.

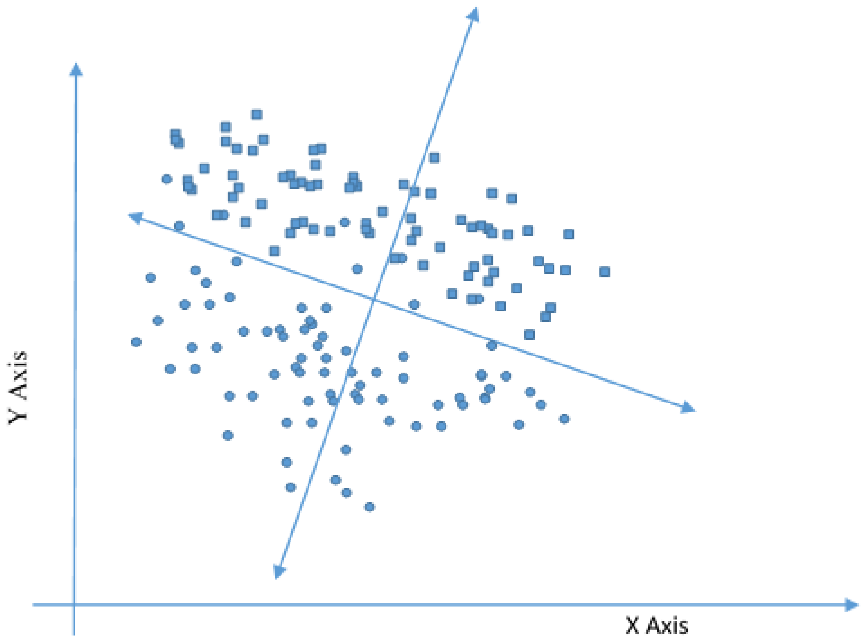

Figure 2.4: Geometrical realisation of PCA

Given a sample of n observation on a vector of p variables $X = (x_1, x_2, x_3, ..., x_p)$. The k^{th} principal component of the sample can be defined as 2.4.0.1.

$$z_k \equiv {a_k}^T X = \sum_{i=1}^{p} a_{ik} x_i \tag{2.4.0.1}$$

Where $a_k = (a_{1k}, a_{2k}, ..., a_{pk})$ is choosen such that $Var[z_k]$ is maximum, subject to $cov[z_k, z_l] = 0$, for $k > k \geq 1$ and to ${a_k}^T a_k = 1$

.

Now to find a_1 first note that:

$$Var[z_1] = {a_1}^T S a_1$$

Where S is the covariance matrix for the variables $X = (x_1, x_2, x_3, ..., x_p)$. As to find a_1, $Var[z_k]$ is maximized subject to ${a_k}^T a_k = 1$, therefore a is an eigenvector of S, corresponding to eigenvalue $\lambda \equiv \lambda_1$. Thus λ_1 is the largest eigen value of S. Hence it can be said that the first principal component has the greatest amount of variation in the sample. So in general the k^{th} principal component can be calculated by equation 2.4.0.2.

$$Var[z_k] = {a_k}^T S a_k = \lambda_k \tag{2.4.0.2}$$

The only problem involving PCA is, it is linear and cannot handle the boundary values properly. Thus as the data comes near to the boundary, the difficulty in classification increases. And towards the last principal component the classfication becomes very vague. To avoid this problem help of support vector machine can be taken.

2.5 Support Vector Machine

Support vector machine can help us avoiding the conflict in the border regions. Support Vector Machines are one among the best off the shelf supervised learning algorithm. They use the optimal margin classifier for classification. In simple words SVM gives an error margin to the classifier line (mathematically known as the regression line). Thus instead of a line as the margin, we have a strip (or hyper-palne) as a classifier. Considering a linear classifier as a binary classification problem with labels y and features x. The notation $y \in \{-1, 1\}$ may be used to denote the class labels. The parameters w & b can be used to write the classifier as given in equation 2.5.0.3. Here $g(z) = 1$ if $z \geq 0$, and $g(z) = -1$ otherwise.

$$h_{w,b}(x) = g(w^T x + b) \tag{2.5.0.3}$$

Given a training exaple $(x^{(i)}, y^{(i)})$, the functional margin of (w, b) with respect to the training example given by equation 2.5.0.4. To increase the functional margin we can rescale w and b, which does not significantly affect any data.

$$\gamma^{(i)} = y^{(i)}(w^T x + b) \tag{2.5.0.4}$$

Talking about geometrical margins, in the figure 2.5 the decision boundary of (w, b) is shown with the w vector which is orthogonal to the separating hyperplane.

The value of $\gamma^{(i)}$ can be found from the equation 2.5.0.5. Finally, given a training set $S = (x^{(i)}, y^{(i)}; i = 1, 2, ..., m$, we also define the geometric margin of (w, b) with respect to S to be the smallest of the geometric margins on the individual training examples. The points with the smallest margins are exactly the ones closest to the decision boundary, which are called the support vectors and they act as the classifiers. [11]

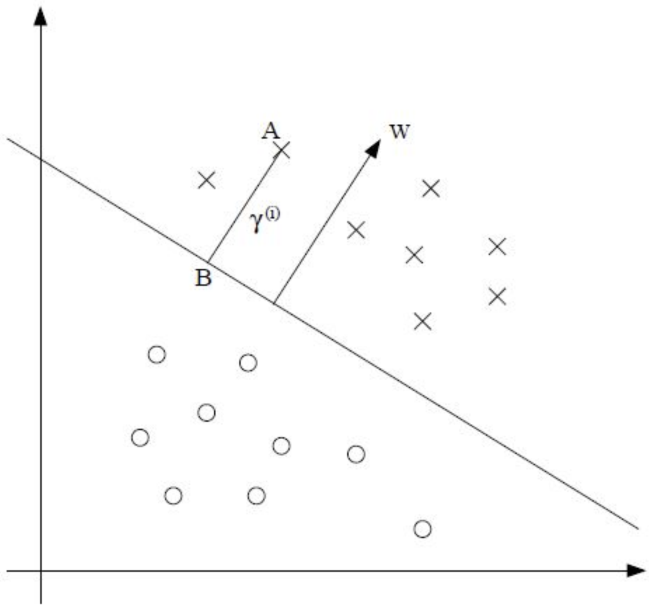

Figure 2.5: Geometrical margin in a SVM

$$\gamma^{(i)} = y^{(i)} \left(\left(\frac{w}{||w||} \right)^T x^{(i)} + \frac{b}{||w||} \right) \tag{2.5.0.5}$$

The SVM can be a linear one or the non-linear one. In this project a simple Non-Linear SVM model is used. What SVM does is to widen the linear classifier to find the support vectors. The SVM is trained and data input then is classified. Here the classifier is trained using the inbuilt matlab command. The non-seizure class is depicted as class 1 (represented as 1 in figure 2.6) and the seizure class is depicted as class 2 (represented as 2 in the figure 2.6). Figure 2.6 represent the classification of seizure and non-seizure data.

Figure 2.6: Comparison of results of seizure and no seizure data using PCA and SVM

2.6 Two Class Non-linear SVM

As a further improvement to the work a two class nonlinear SVM model was implemented.The following gives the pseudo code algorithm for implementation of the SVM model.

```
Set matrix x = [Test sample];
Matrix X = [Training samples];
Matrix Y = [Class identify];
Matrix B = Y';
for i = 1 to n (n-No of test samples)
  for j = 1 to n
     Calculate classification function parameters using matrix X and Gaussian radial basis
     function and give the result to matrix
     A(i,j) = exp[ -((X(i,1) - X(j,1))^2 + ...... + (X(i,39) - X(j,39))^2 / 2σ^2 ];
  end
end
for k = 1 to n
  for l = 1 to n
     A(k,l) = A(k,l) × Y(l);
  end
end
```

```
C = A^-1 × B;
for q = 1to n
   h(q) = C(q) × Y(q) × exp[ -((X(q,1) - X(1,1))^2 + ..... + (X(q,39) - X(1,39))^2 / 2σ^2 ];
end
H = Sum of all elements in matrix h;
b = Y(1) - H;
for p = 1 to n
   g(p) = C(p) × Y(p) × exp[ -((X(p,1) - x(1))^2 + ..... + (X(p,39) - x(39))^2 / 2σ^2 ];
end
W = Sum of all elements in matrix g;
Build the classification functions G = W + b;
```

$$\begin{cases} \operatorname{sgn}\{G\} = 1, X \in C1 \\ \operatorname{sgn}\{G\} = -1, X \in C2 \end{cases}$$

The algorithm was applied to the IRIS data set. The following table gives the results of the classification.

	training data	testng data	Result SVM	Error SVM	Result Code	Error Code
Iris-Setosa	30	20	20	0	20	0
Iris-Versicolor	30	20	20	0	20	0
	training data	testng data	Result SVM	Error SVM	Result Code	Error Code
Iris-Versicolor	30	20	19	1	19	1
Iris-Verginica	30	20	19	1	18	2
	training data	testng data	Result SVM	Error SVM	Result Code	Error Code
Iris-Setosa	30	20	20	0	20	0
Iris-Verginica	30	20	20	0	20	0

Table 2.1: Test results for IRIS Data set

As the classifier has started showing promising results, the algorithm was ready to be implemented for emotion recognition. All the algorithms tested for epilepsy was further applied for emotion detection, which is discussed in the next chapter.

Chapter 3

EMOTION DETECTION

In cognitive science it is believed that human emotion is not such a thing that can be quantified or generalised, because the extent of emotions are different for different persons. Thus for any algorithm to work it is necessary that it should be tested on persons having almost similar emotional states and responses. A study conducted in the Queen Marry University, UK collected data for several subjects and tried to make a database for emotion detection - DEAP (A Database for Emotion Analysis using Physiological Signals)[4]. The database is open source, but with an agreement of using the data for study purpose only. A valid experiment was conducted there to find the emotional response of the subjects.

3.1 The Experiment

32 Healthy participants (50 percent female,), aged between 19 and 37 (mean age 26.9), took part in the experiment. Before the experiment, each participant was asked to sign a consent form and fill out a questionnaire. After that, they were given a set of instructions to read, informing them of the experiment rules and regulations and the definition of the different scales used for self-assessment. An experimenter was there who was responsible to answer any questions. As the participant got a clear picture of the instructions he moved on for the experiment.

The participants were shown music video extracts of having the potential to arouse a particular kind of emotion in each of the participants. Each one of them was shown 40 videos. These videos were selected by volunteers. Then the data was taken from 40 channel EEG and first sampled at 512 Hz and then down-sampled to 128Hz. After the subjects were subjected to the test, and their signals were recorded, then another response paper was given to them to rate their emotional status. The emotions were classified into 4 different classes as arousal, dominance,liking and valence. Arousal means from inactive (e.g. uninterested, bored) to active (e.g. alert, excited), whereas valence means from unpleasant (e.g.sad, stressed) to pleasant (e.g. happy, elated). Dominance is something that ranges from a helpless and weak feeling (without control) to an empowered feeling (in control of everything). [4]

3.2 Classification

Further the data and features were given to a learning algorithm. A simple linear classifier was used to classify the data. The SVM was trained using the inbuilt function of matlab. Figure 3.1, 3.2, 3.3, 3.4, 3.5 shows the results of the emotion classification using the modified SVM approach. It is clear from the figures that with advanced standard algorithms it is not impossible to classify and detect emotions.

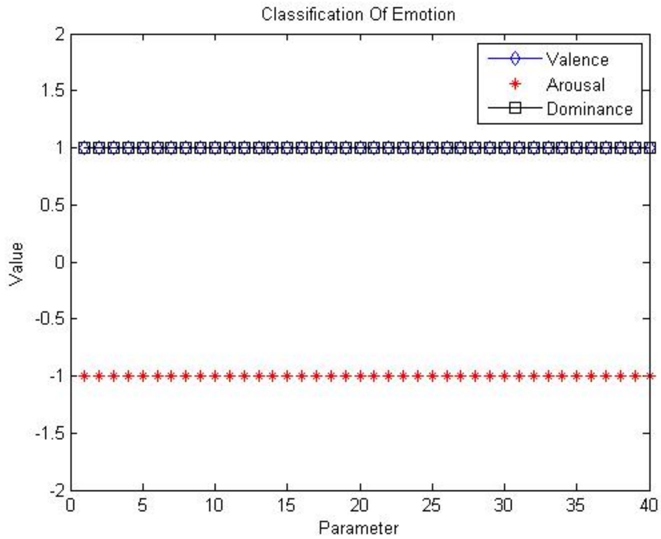

Figure 3.1: Emotion data classification - Person1 Video 5

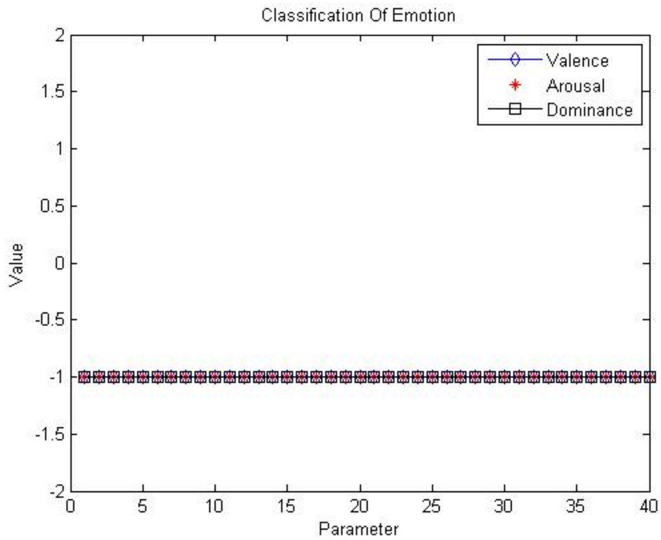

Figure 3.2: Emotion data classification - Person1 Video 9

Further for analysis purpose the algorithm discussed in 2.6 is applied here for emotion detection. The data collected were tabulated and analysed.

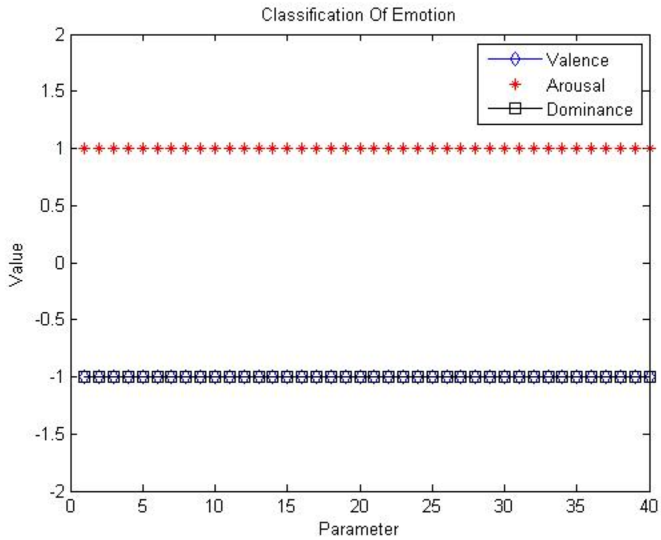

Figure 3.3: Emotion data classification - Person1 Video 15

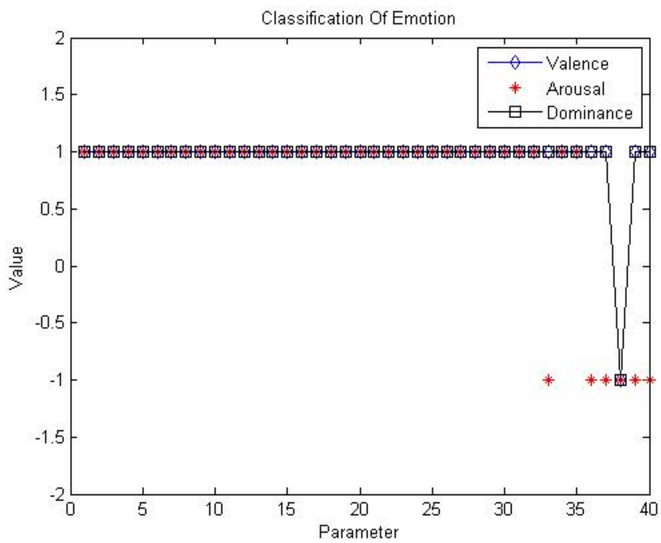

Figure 3.4: Emotion data classification - Person1 Video 21

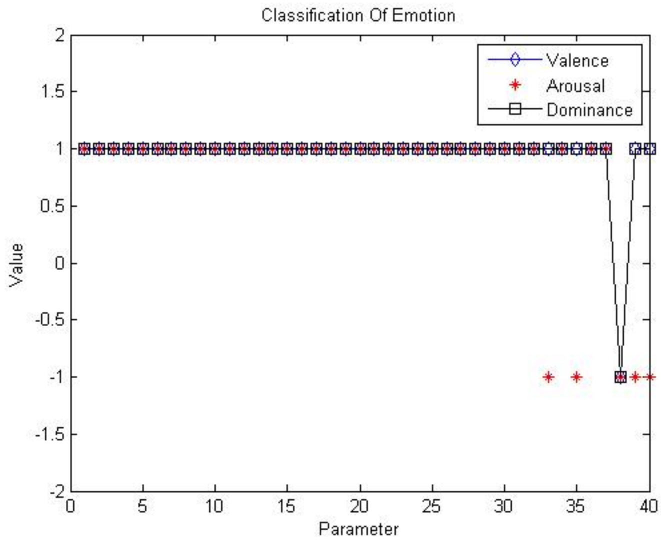

Figure 3.5: Emotion data classification - Person1 Video 23

	a.Label	Balance	SVM.Balance	SVM.Error	Code.Balance	Code.Error
Per1_vid21	9.00	1	1	No	1	No
Per1_vid22	7.09	1	1	No	-1	Yes
Per1_vid23	8.15	1	1	No	-1	Yes
Per1_vid24	7.04	1	1	No	-1	Yes
Per1_vid25	8.86	1	1	No	1	No
Per1_vid26	7.28	1	1	No	1	No
Per1_vid27	7.35	1	1	No	1	No
Per1_vid28	3.88	-1	1	Yes	-1	No
Per1_vid29	1.36	-1	-1	No	-1	No
Per1_vid30	2.08	-1	1	Yes	1	Yes
Per1_vid31	3.03	-1	1	Yes	-1	No
Per1_vid32	2.28	-1	1	Yes	-1	No
Per1_vid33	3.81	-1	1	Yes	-1	No
Per1_vid34	2.28	-1	1	Yes	-1	No
Per1_vid35	2.06	-1	-1	No	-1	No
Per1_vid36	2.90	-1	1	Yes	-1	No
Per1_vid37	2.31	-1	-1	No	-1	No
Per1_vid38	3.33	-1	1	Yes	-1	No
Per1_vid39	3.24	-1	1	Yes	-1	No
Per1_vid40	5.10	1	1	No	-1	Yes

Table 3.1: Test results for emotion data set: Total SVM error = 9, total code error = 5

Chapter 4

A STEP TOWARDS HARDWARE IMPLEMENTATIONS

The goal of the project was to take a new step towards the emotion detection. All the previous chapters dealt with the detection of emotion using softwares. The figure 4.1 shows the flow diagram of the system architecture.The previous chapters covered classification using inbuilt and user defined functions. So after checking the accuracy of the code, the simplified Non-Linear SVM code was written in VHDL . The next stage describes briefly about the hardware realisation.

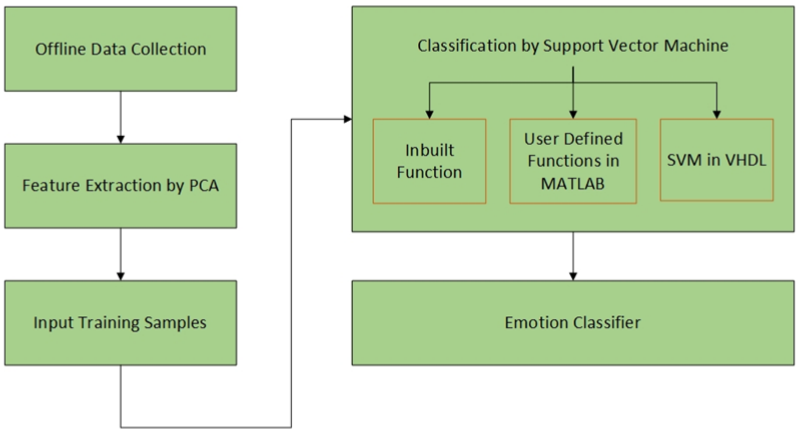

Figure 4.1: Flow diagram of the architecture

VHDL (VHSIC Hardware Description Language) is commonly used to translate a logical program into its corresponding hardware architecture. Before designing any code in VHDL a prototype architecture was designed, a block diagram (figure) representing all the blocks needed for converting the idea of emotion recognition into a hardware architecture.

The VHDL code was developed in XILINX ISE and this chapter gives the corresponding output waveform of all the blocks. The challenge faced during this was developing a hardware module for calculating the inverse of a matrix, which is still a research area. Thus a software interface will continuously be running with the system to take the output of the hardware and convert it to the inverse matrix of the input data. The following figures give the output of the VHDL codes and testbenches.

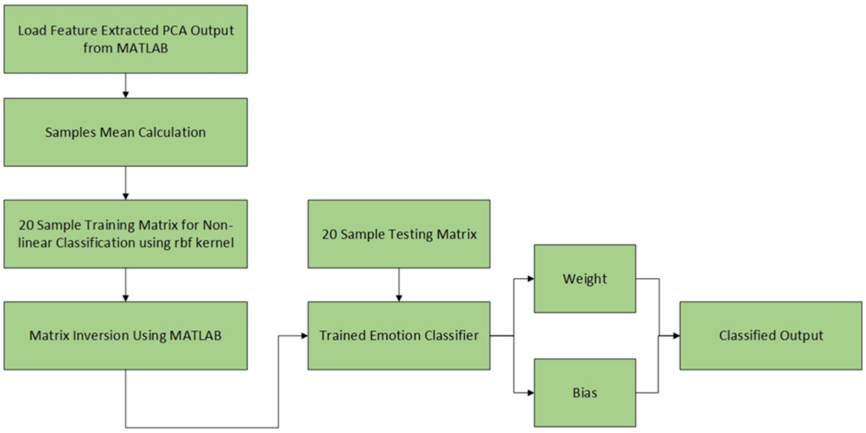

Figure 4.2: Flow diagram of the prototyped VHDL system architecture

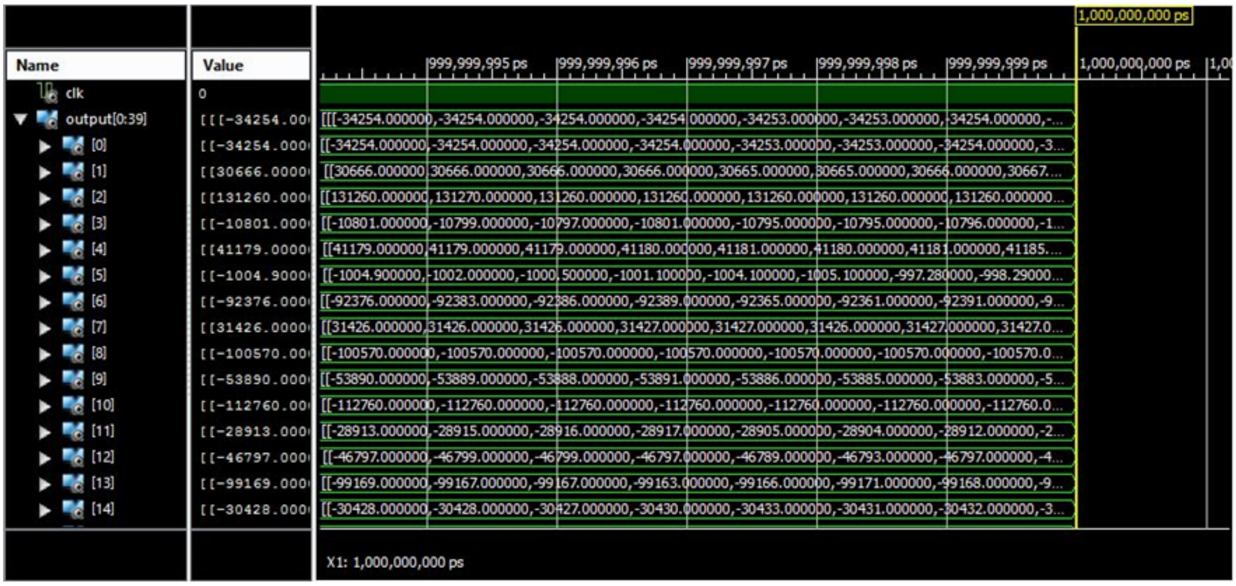

Figure 4.3: Loaded Files From Matlab

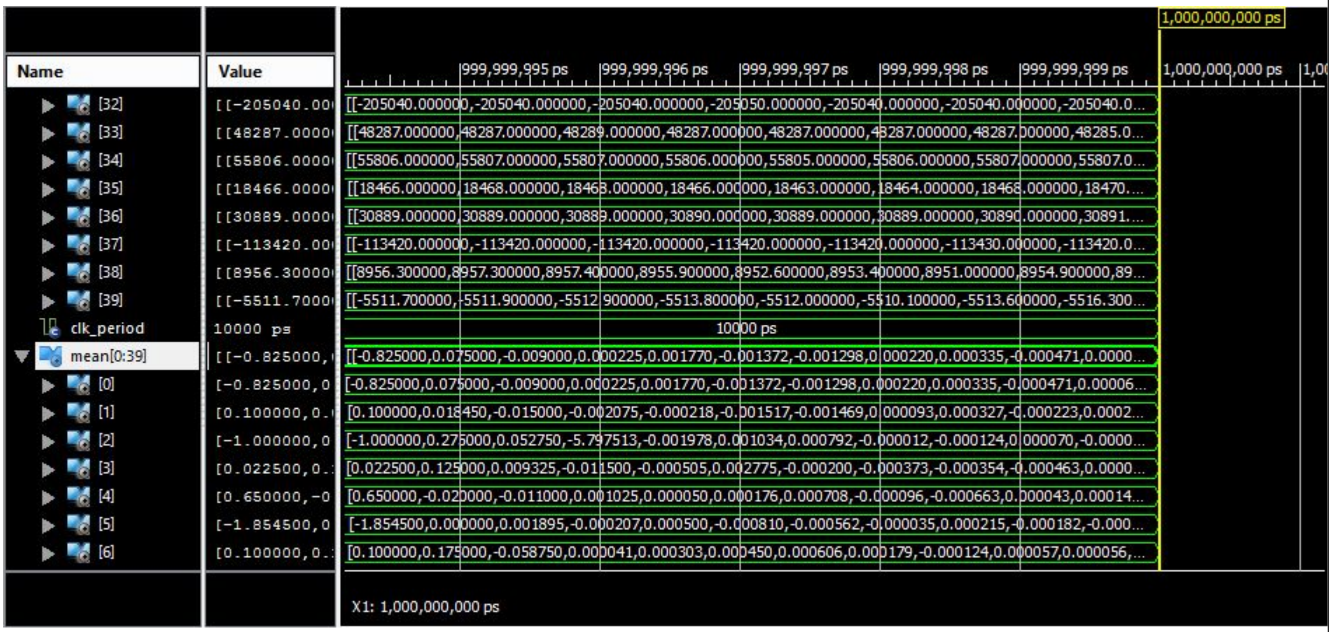

Figure 4.4: Mean Of Each Sample

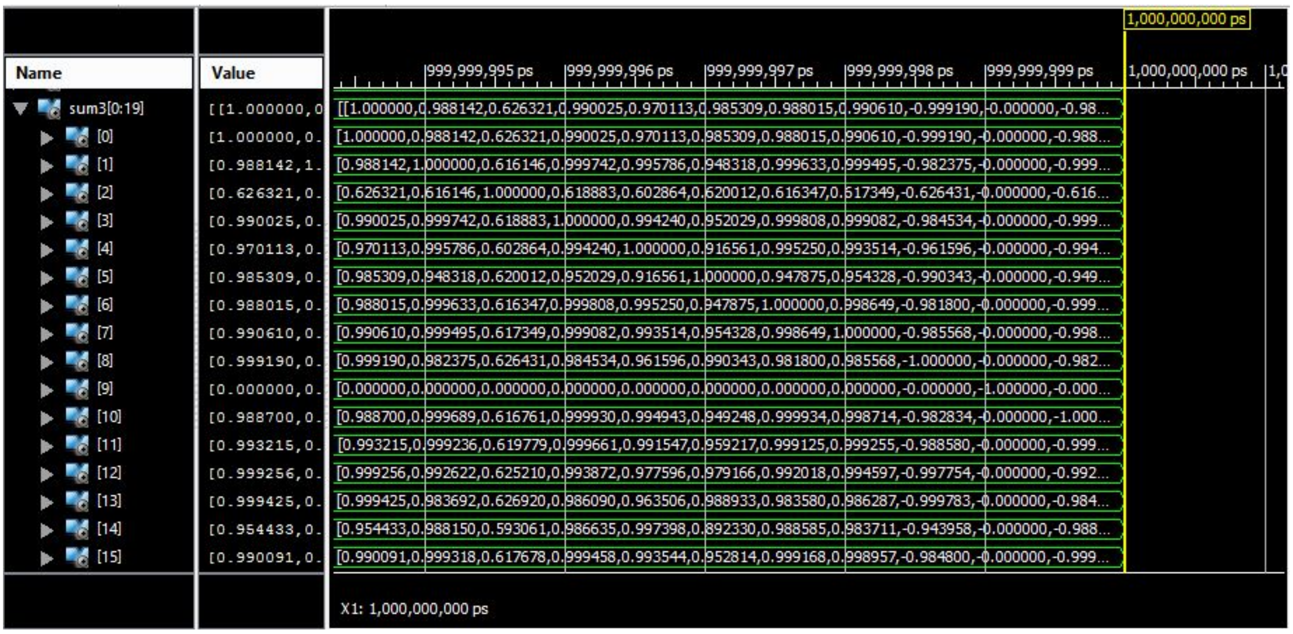

Figure 4.5: Matrix A-For Training

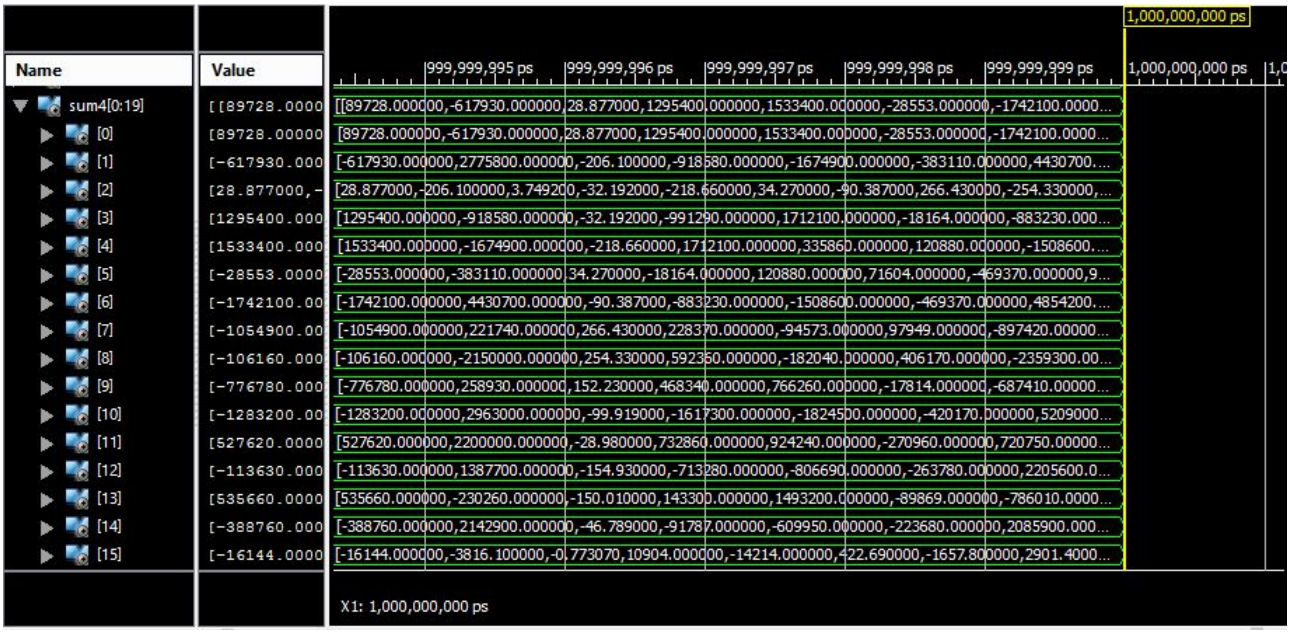

Figure 4.6: Inverse of Matrix A

Figure 4.7: Training Matrix

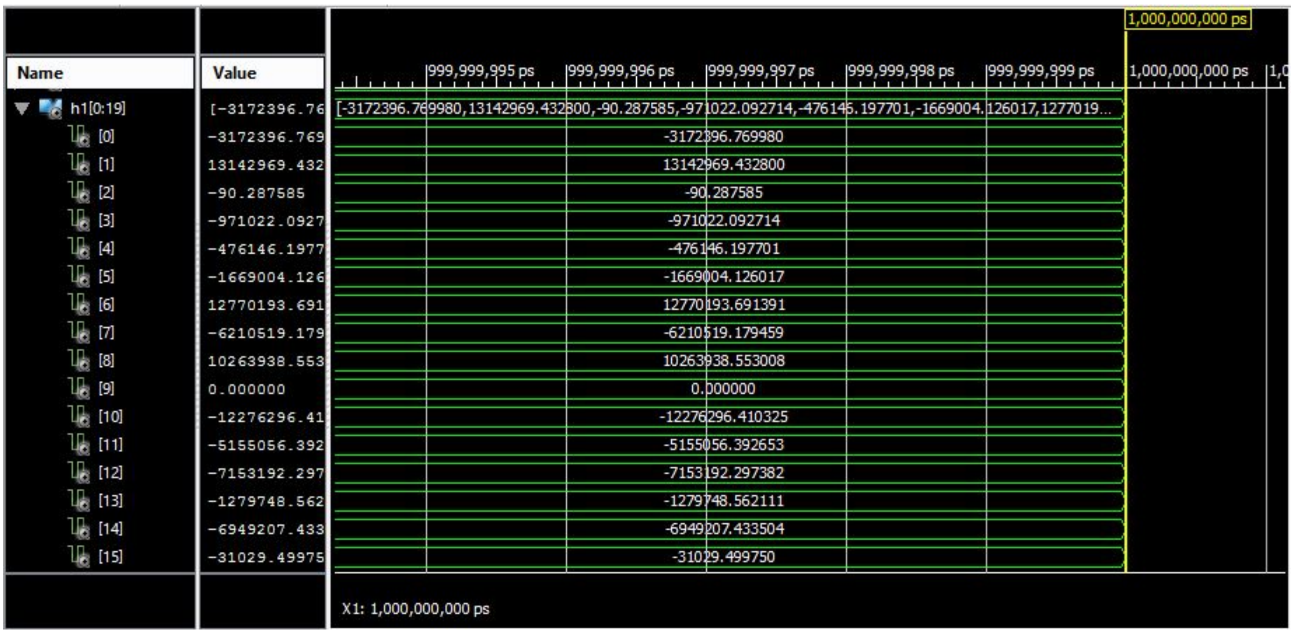

Figure 4.8: Evaluation of bias term

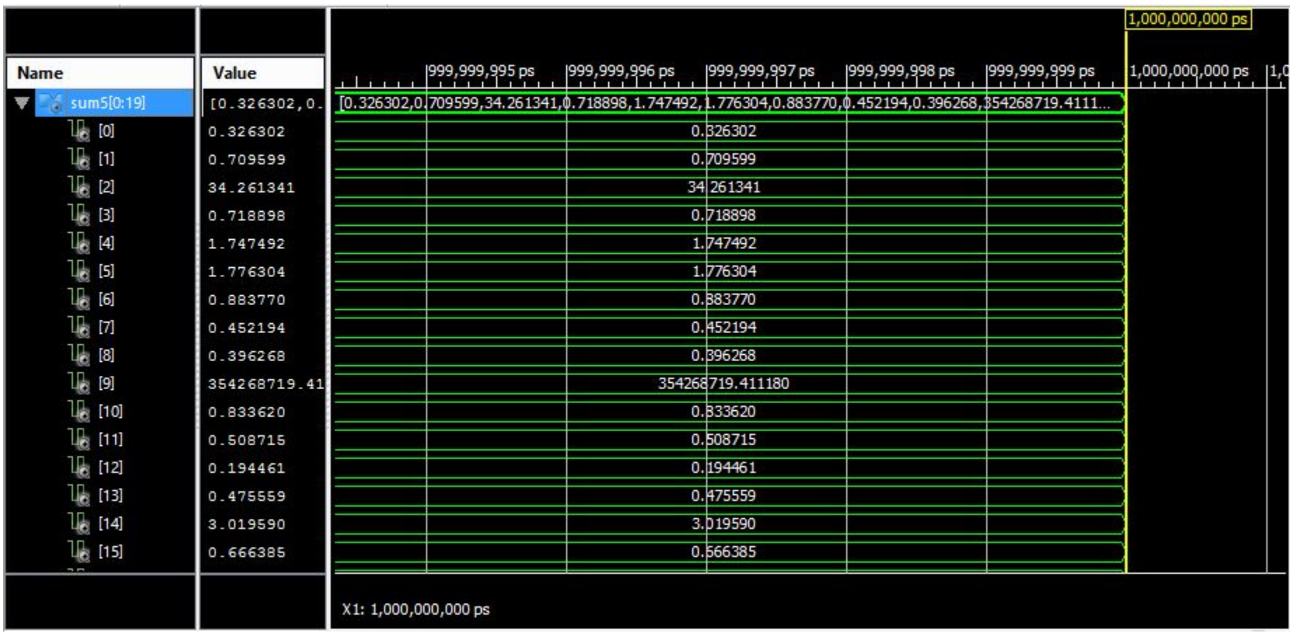

Figure 4.9: Testing on Input Sample

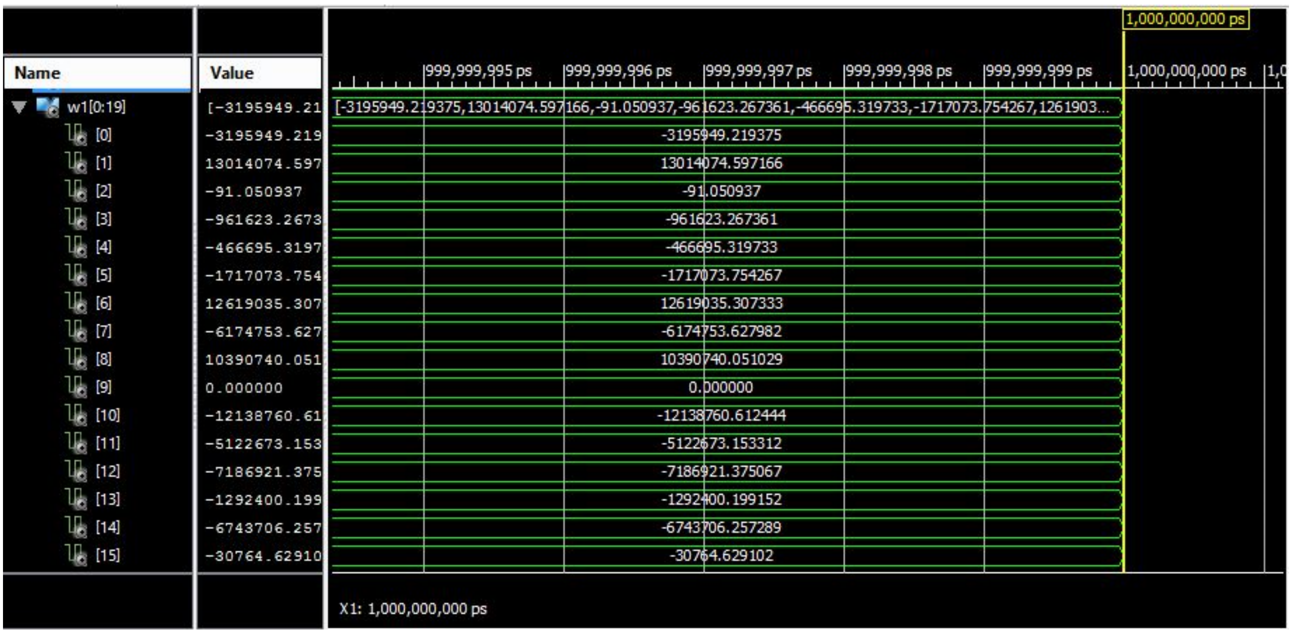

Figure 4.10: Evaluation of weight

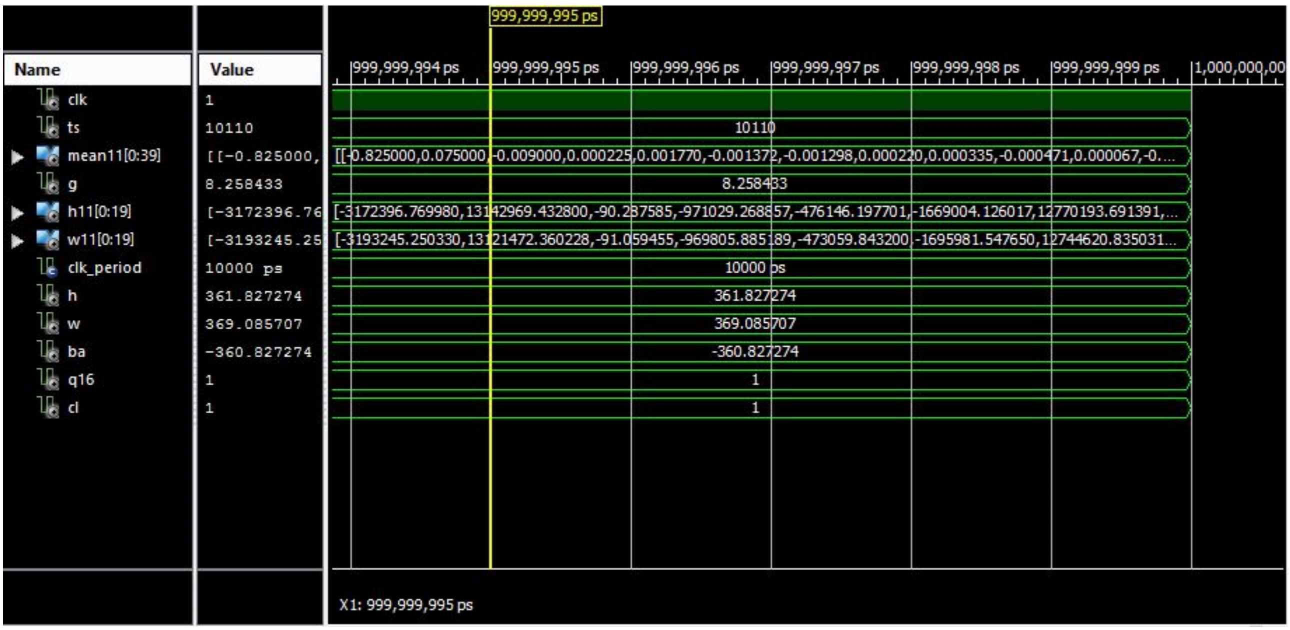

Figure 4.11: Output For Vid 23-Positive Valence

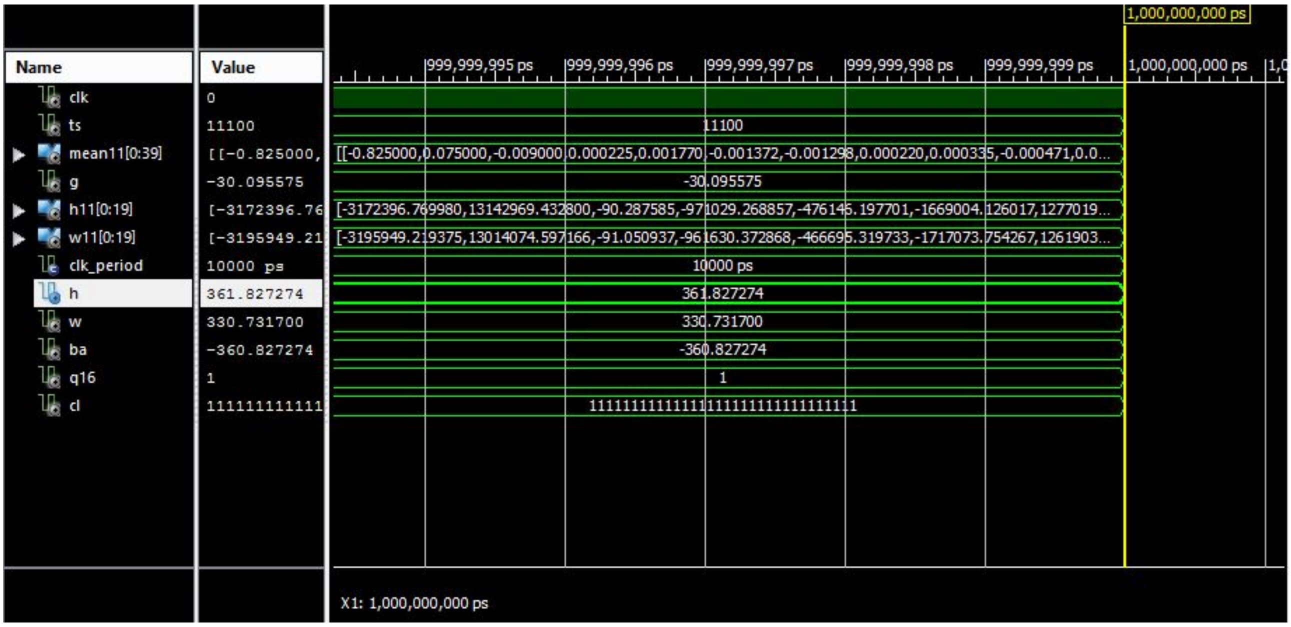

Figure 4.12: Output For Vid 29-Negative Valence

Chapter 5

CONCLUSION

The proposed algorithm gave significant results for the detection of the specific emotions. The linearity of the classifier went through fair amount of speed breakers, still giving almost proper results. This classifier was working with high perfection for detection of epilepsy of a patient. But significant study is required to generalize the algorithm as there is no proper bench mark for the brain activities. Further improvisation can be done by using ICA instead of PCA and using a multiclass SVM.

With the nonlinear SVM, the output efficiency was increased and the most important step was thinking of implementing this in hardware. Given this is a signal processing field, the software code can be dumped into a DSP and the system could work just fine. But a hardware architecture will give both portability and abstraction to the design. The written VHDL codes were not synthesizable. Thus a proper RTL implementation is necessary. As the code deals with a large amount of data it will have to be interfaced with a very large memory. A later stage analysis using cadence or synopsis will give us more knowledge on power consumption. But yet a synthesizable code remains a challenge for the future scope of the project. If developed properly over a longterm research, it can help us produce an on chip emotion detector.

Bibliography

[1] Panagiotis C. Petrantonakis, Leontios J. Hadjileontiadis, ”A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG-based emotion recognition,” IEEE Trans. on Information Tech. and bio-medicine., vol. 15, no. 5, pp. 737-746,2011.

[2] M. Murugappan, M. Rizon, R. Nagarajan, and S. Yaacob, ”Inferring of human emotional states using multichannel EEG”, European Journal of Scientific Research ISSN 1450-216X vol.48 No.2, pp. 281-299, 2010.

[3] A. S. AlMejrad, ”Human emotions detection using brain wave signals: a challenging.” European Journal of Scientific Research, vol. 44, No. 4, pp. 640-659, 2010.

[4] S. Koelstra, C. Muehl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras, ”DEAP: A Database for Emotion Analysis using Physiological Signals”, 2013, ”http://www.eecs.qmul.ac.uk/mmv/datasets/deap/index.html”

[5] Ali Shoeb, ”Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment”, 2000, ”http://www.physionet.org/cgi-bin/atm/ATM”

[6] P. Ekman and W. V. Friesen ”Unmasking the Face: A Guide to Recognizing Emotions From Facial Clues,” Englewood Cliffs, NJ:Prentice-Hall, 1975.

[7] H. Kobayashi and F. Hara : ”The recognition of basic facial expressions by neural network,” Trans. Soc. Instrum. Contr. Eng, vol. 29, no. 1, pp.112-118,1993.

[8] Anisha Halder, Rajshree MandaI, Pavel Bhowmik, Aruna Chakraborty, Amit Konar and Nikhil R. Pal, ”General and Interval Type-2 Fuzzy Face-Space Approach to Emotion Recognition,” IEEE Trans. On Systems, Man and Cybernetics, Part-A, vol. 43, no. 3, pp. 587-605, May, 2013.

[9] J.C. Platt: A Fast Algorithm for Training Support Vector Machines, Available: http://research.microsoft.com/users/jplatt/smo.html

[10] R.-E. Fan, P.-H. Chen, and C.-J. Lin. Working Set Selection Using Second Order Information for Training SVM. Journal of Machine Learning Research, 6(2005), 1889-1918.

[11] L. Bottou and C.-J. Lin. Support Vector Machine Solvers. 2006, Available: http://www.csie.ntu.edu.tw/ cjlin/papers.html

[12] Xiaohui Song , “FPGA Implementation of a Support Vector Machine based Classification System and its Potential Application in Smart Grid.”, M.S. Thesis, Department of Electrical Engineering, The University of Toledo, December 2013